

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/572,976
Source: JFWP
Date Processed by STIC: 04/10/2006

ENTERED

IFWP

RAW SEQUENCE LISTING

DATE: 04/10/2006

PATENT APPLICATION: US/10/572,976

TIME: 10:44:52

Input Set : F:\6550-086POA SeqListing.ST25.txt
 Output Set: N:\CRF4\04102006\J572976.raw

3 <110> APPLICANT: Board of Trustees operating Michigan State University
 4 Frost, John W.
 6 <120> TITLE OF INVENTION: Methods and Materials for the Production of Shikimic Acid
 8 <130> FILE REFERENCE: 6550-000086
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/572,976
 C--> 10 <141> CURRENT FILING DATE: 2006-03-22
 10 <150> PRIOR APPLICATION NUMBER: US 60/505,658
 11 <151> PRIOR FILING DATE: 2003-09-24
 13 <160> NUMBER OF SEQ ID NOS: 34
 15 <170> SOFTWARE: PatentIn version 3.3
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 618
 19 <212> TYPE: DNA
 20 <213> ORGANISM: Escherichia coli
 23 <220> FEATURE:
 24 <221> NAME/KEY: CDS
 25 <222> LOCATION: (1)..(615)
 26 <223> OTHER INFORMATION: dgoA CDS for KDPGal Aldolase
 28 <400> SEQUENCE: 1
 29 atg cag tgg caa act aaa ctc ccg ctg atc gcc att ttg cgc ggt att 48
 30 Met Gln Trp Gln Thr Lys Leu Pro Leu Ile Ala Ile Leu Arg Gly Ile
 31 1 5 10 15
 33 acg ccc gac gag gcg ctg gcg cat gtt ggc gcg gtg att gac gcc ggg 96
 34 Thr Pro Asp Glu Ala Leu Ala His Val Gly Ala Val Ile Asp Ala Gly
 35 20 25 30
 37 ttc gac gcg gtt gaa atc ccg ctg aat tcc cca caa tgg gag caa agc 144
 38 Phe Asp Ala Val Glu Ile Pro Leu Asn Ser Pro Gln Trp Glu Gln Ser
 39 35 40 45
 41 att ccc gcc atc gtt gat gcg tac ggc gac aag gcg ttg att ggc gca 192
 42 Ile Pro Ala Ile Val Asp Ala Tyr Gly Asp Lys Ala Leu Ile Gly Ala
 43 50 55 60
 45 ggt acg gta ctg aaa cct gaa cag gtc gat gcg ctc gcc agg atg ggc 240
 46 Gly Thr Val Leu Lys Pro Glu Gln Val Asp Ala Leu Ala Arg Met Gly
 47 65 70 75 80
 49 tgt cag ctc atc gtt acg ccc aat atc cat agt gaa gtg atc cgc cgt 288
 50 Cys Gln Leu Ile Val Thr Pro Asn Ile His Ser Glu Val Ile Arg Arg
 51 85 90 95
 53 gcg gtg ggc tac ggc atg acc gtc tgc ccc ggc tgc gcg acg gcg acc 336
 54 Ala Val Gly Tyr Gly Met Thr Val Cys Pro Gly Cys Ala Thr Ala Thr
 55 100 105 110
 57 gaa gcc ttt acc gcg ctc gaa gcg ggc gcg cag gcg ctg aaa ata ttt 384
 58 Glu Ala Phe Thr Ala Leu Glu Ala Gly Ala Gln Ala Leu Lys Ile Phe
 59 115 120 125

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/572,976

DATE: 04/10/2006
TIME: 10:44:52

Input Set : F:\6550-086POA SeqListing.ST25.txt
Output Set: N:\CRF4\04102006\J572976.raw

61	ccg tca tcg gct ttt ggt ccg caa tac atc aaa	gcg tta aaa	gcg gta	432
62	Pro Ser Ser Ala Phe Gly Pro Gln Tyr Ile Lys	Ala Leu Lys	Ala Val	
63	130	135	140	
65	ttg cca tcg gac atc gca gtc ttt gcc gtt ggc	ggc gtg acg	cca gaa	480
66	Leu Pro Ser Asp Ile Ala Val Phe Ala Val	Gly Gly Val	Thr Pro Glu	
67	145	150	155	160
69	aac ctg gcg cag tgg ata gac gca ggt	tgt gca	ggg gcg ggc tta	528
70	70 Asn Leu Ala Gln Trp Ile Asp Ala Gly Cys	Ala Gly	Leu Gly	
71	165	170	175	
73	73 agc gat ctc tat cgc gcc ggg caa tcc gta	gag cgc acc	gcg cag cag	576
74	Ser Asp Leu Tyr Arg Ala Gly Gln Ser Val	Glu Arg Thr	Ala Gln Gln	
75	180	185	190	
77	77 gca gca gca ttt gtt aag gcg tat cga	gag gca	gtg caa tga	618
78	78 Ala Ala Ala Phe Val Lys Ala Tyr Arg	Glu Ala Val	Gln	
79	195	200	205	
82	<210> SEQ ID NO: 2			
83	<211> LENGTH: 205			
84	<212> TYPE: PRT			
85	<213> ORGANISM: Escherichia coli			
87	<400> SEQUENCE: 2			
89	89 Met Gln Trp Gln Thr Lys Leu Pro Leu Ile	Ala Ile Leu Arg	Gly Ile	
90	1	5	10	15
93	93 Thr Pro Asp Glu Ala Leu Ala His Val	Gly Ala Val	Ile Asp Ala Gly	
94	20	25	30	
97	97 Phe Asp Ala Val Glu Ile Pro Leu Asn Ser	Pro Gln Trp Glu Gln Ser		
98	35	40	45	
101	101 Ile Pro Ala Ile Val Asp Ala Tyr Gly	Asp Lys Ala Leu Ile Gly	Ala	
102	50	55	60	
105	105 Gly Thr Val Leu Lys Pro Glu Gln Val	Asp Ala Leu Ala Arg	Met Gly	
106	65	70	75	80
109	109 Cys Gln Leu Ile Val Thr Pro Asn Ile	His Ser Glu Val Ile Arg	Arg	
110	85	90	95	
113	113 Ala Val Gly Tyr Gly Met Thr Val Cys	Pro Gly Cys Ala Thr	Ala Thr	
114	100	105	110	
117	117 Glu Ala Phe Thr Ala Leu Glu Ala Gly	Ala Gln Ala Leu Lys	Ile Phe	
118	115	120	125	
121	121 Pro Ser Ser Ala Phe Gly Pro Gln Tyr	Ile Lys Ala Leu Lys	Ala Val	
122	130	135	140	
125	125 Leu Pro Ser Asp Ile Ala Val Phe Ala	Val Gly Gly Val	Thr Pro Glu	
126	145	150	155	160
129	129 Asn Leu Ala Gln Trp Ile Asp Ala Gly	Cys Ala Gly Ala Gly	Leu Gly	
130	165	170	175	
133	133 Ser Asp Leu Tyr Arg Ala Gly Gln Ser	Val Glu Arg Thr	Ala Gln Gln	
134	180	185	190	
137	137 Ala Ala Ala Phe Val Lys Ala Tyr Arg	Glu Ala Val Gln		
138	195	200	205	
141	<210> SEQ ID NO: 3			
142	<211> LENGTH: 618			
143	<212> TYPE: DNA			

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/572,976

DATE: 04/10/2006

TIME: 10:44:52

Input Set : F:\6550-086POA SeqListing.ST25.txt
 Output Set: N:\CRF4\04102006\J572976.raw

144 <213> ORGANISM: Klebsiella pneumoniae
 147 <220> FEATURE:
 148 <221> NAME/KEY: CDS
 149 <222> LOCATION: (1)..(615)
 150 <223> OTHER INFORMATION: dgoA CDS for KDPGal Aldolase
 152 <400> SEQUENCE: 3

153 atg cag tgg caa act aac ctt cca ctt atc gct atc ctg cgc ggt att	48
154 Met Gln Trp Gln Thr Asn Leu Pro Leu Ile Ala Ile Leu Arg Gly Ile	
155 1 5 10 15	
157 acg cca gac gag gcg ctg gct cac gtt gcc gtc gtt atc gac gcc ggt	96
158 Thr Pro Asp Glu Ala Leu Ala His Val Gly Ala Val Ile Asp Ala Gly	
159 20 25 30	
161 ttc gac gcg gtc gaa atc ccg ctg aac tcg ccg cag tgg gag aaa agt	144
162 Phe Asp Ala Val Glu Ile Pro Leu Asn Ser Pro Gln Trp Glu Lys Ser	
163 35 40 45	
165 att ccg cag gtc gac gct tac ggc gag cag gcg ctt atc ggc gcg	192
166 Ile Pro Gln Val Val Asp Ala Tyr Gly Glu Gln Ala Leu Ile Gly Ala	
167 50 55 60	
169 ggc acg gtg ctg caa ccg gag cag gtc gac agg ctg gcg gcc atg ggc	240
170 Gly Thr Val Leu Gln Pro Glu Gln Val Asp Arg Leu Ala Ala Met Gly	
171 65 70 75 80	
173 tgt ccg ctg att gtg acg cca aac att cca ccg gaa gtg atc ccg cga	288
174 Cys Arg Leu Ile Val Thr Pro Asn Ile Gln Pro Glu Val Ile Arg Arg	
175 85 90 95	
177 gcg gtg ggt tac ggc atg acc gtg tgt cca ggc tgc gcc acc gcc agc	336
178 Ala Val Gly Tyr Gly Met Thr Val Cys Pro Gly Cys Ala Thr Ala Ser	
179 100 105 110	
181 gaa gcc ttt agc gcg ctc gat gcc ggc gcg cag gcg cta aaa atc ttc	384
182 Glu Ala Phe Ser Ala Leu Asp Ala Gly Ala Gln Ala Leu Lys Ile Phe	
183 115 120 125	
185 ccg tca tcg gct ttt ggc ccg gat tac atc aaa gcg ttg aaa gcc gtg	432
186 Pro Ser Ser Ala Phe Gly Pro Asp Tyr Ile Lys Ala Leu Lys Ala Val	
187 130 135 140	
189 ctg ccg ccc gag gtt ccg gtc ttt gcc gtt ggc ggc gtg acg ccg gaa	480
190 Leu Pro Pro Glu Val Pro Val Phe Ala Val Gly Gly Val Thr Pro Glu	
191 145 150 155 160	
193 aac ctg gcg cag tgg att aat gcc ggc tgt gtt ggg gca gga ttg ggt	528
194 Asn Leu Ala Gln Trp Ile Asn Ala Gly Cys Val Gly Ala Gly Leu Gly	
195 165 170 175	
197 agc gat ctc tat cgt gcc ggc cag tcg gtt gaa cgt acc gcg cag cag	576
198 Ser Asp Leu Tyr Arg Ala Gly Gln Ser Val Glu Arg Thr Ala Gln Gln	
199 180 185 190	
201 gca gcc gca ttc gta aaa gcg tat cga gag gca gtg aaa tga	618
202 Ala Ala Ala Phe Val Lys Ala Tyr Arg Glu Ala Val Lys	
203 195 200 205	
206 <210> SEQ ID NO: 4	
207 <211> LENGTH: 205	
208 <212> TYPE: PRT	
209 <213> ORGANISM: Klebsiella pneumoniae	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/572,976

DATE: 04/10/2006
TIME: 10:44:52

Input Set : F:\6550-086POA SeqListing.ST25.txt
Output Set: N:\CRF4\04102006\J572976.raw

211 <400> SEQUENCE: 4
 213 Met Gln Trp Gln Thr Asn Leu Pro Leu Ile Ala Ile Leu Arg Gly Ile
 214 1 5 10 15
 217 Thr Pro Asp Glu Ala Leu Ala His Val Gly Ala Val Ile Asp Ala Gly
 218 20 25 30
 221 Phe Asp Ala Val Glu Ile Pro Leu Asn Ser Pro Gln Trp Glu Lys Ser
 222 35 40 45
 225 Ile Pro Gln Val Val Asp Ala Tyr Gly Glu Gln Ala Leu Ile Gly Ala
 226 50 55 60
 229 Gly Thr Val Leu Gln Pro Glu Gln Val Asp Arg Leu Ala Ala Met Gly
 230 65 70 75 80
 233 Cys Arg Leu Ile Val Thr Pro Asn Ile Gln Pro Glu Val Ile Arg Arg
 234 85 90 95
 237 Ala Val Gly Tyr Gly Met Thr Val Cys Pro Gly Cys Ala Thr Ala Ser
 238 100 105 110
 241 Glu Ala Phe Ser Ala Leu Asp Ala Gly Ala Gln Ala Leu Lys Ile Phe
 242 115 120 125
 245 Pro Ser Ser Ala Phe Gly Pro Asp Tyr Ile Lys Ala Leu Lys Ala Val
 246 130 135 140
 249 Leu Pro Pro Glu Val Pro Val Phe Ala Val Gly Gly Val Thr Pro Glu
 250 145 150 155 160
 253 Asn Leu Ala Gln Trp Ile Asn Ala Gly Cys Val Gly Ala Gly Leu Gly
 254 165 170 175
 257 Ser Asp Leu Tyr Arg Ala Gly Gln Ser Val Glu Arg Thr Ala Gln Gln
 258 180 185 190
 261 Ala Ala Ala Phe Val Lys Ala Tyr Arg Glu Ala Val Lys
 262 195 200 205
 265 <210> SEQ ID NO: 5
 266 <211> LENGTH: 618
 267 <212> TYPE: DNA
 268 <213> ORGANISM: Salmonella typhimurium
 271 <220> FEATURE:
 272 <221> NAME/KEY: CDS
 273 <222> LOCATION: (1)..(615)
 274 <223> OTHER INFORMATION: dgoA CDS for KDPGal Aldolase
 276 <400> SEQUENCE: 5
 277 atg cag tgg caa act aat ctc cct ctc atc gct atc tta cgc ggt att 48
 278 Met Gln Trp Gln Thr Asn Leu Pro Leu Ile Ala Ile Leu Arg Gly Ile
 279 1 5 10 15
 281 acg ccc gat gat gcc ctg gcg cac gtt ggc gcg gtg gtg gat gcg gga 96
 282 Thr Pro Asp Asp Ala Leu Ala His Val Gly Ala Val Val Asp Ala Gly
 283 20 25 30
 285 ttt gac gct ata gaa att ccg ctt aac tcc cca cag tgg gaa aaa agc 144
 286 Phe Asp Ala Ile Glu Ile Pro Leu Asn Ser Pro Gln Trp Glu Lys Ser
 287 35 40 45
 289 att tct tcc gtg gtg aag gcg tat ggc ggc agg gcg ctt att ggc gct 192
 290 Ile Ser Ser Val Val Lys Ala Tyr Gly Gly Arg Ala Leu Ile Gly Ala
 291 50 55 60
 293 ggt acc gta ctg aaa ccg gaa cag gta gac cag ctt gcc ggg atg ggc 240

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/572,976

DATE: 04/10/2006

TIME: 10:44:52

Input Set : F:\6550-086POA SeqListing.ST25.txt
 Output Set: N:\CRF4\04102006\J572976.raw

294	Gly	Thr	Val	Leu	Lys	Pro	Glu	Gln	Val	Asp	Gln	Leu	Ala	Gly	Met	Gly
295	65						70				75				80	
297	tgc	aag	ctg	atc	gtc	acg	ccg	aat	atc	caa	ccg	gag	gtg	atc	cgc	cg
298	Cys	Lys	Leu	Ile	Val	Thr	Pro	Asn	Ile	Gln	Pro	Glu	Val	Ile	Arg	Arg
299															95	
301	gca	gtg	agc	tat	ggc	atg	acc	gtg	tgt	ccg	ggc	tgc	gcc	acg	gca	acg
302	Ala	Val	Ser	Tyr	Gly	Met	Thr	Val	Cys	Pro	Gly	Cys	Ala	Thr	Ala	Thr
303															110	
305	gaa	gcc	ttt	tct	gca	ctg	gat	gca	ggc	gca	cag	gca	tta	aaa	att	ttc
306	Glu	Ala	Phe	Ser	Ala	Leu	Asp	Ala	Gly	Ala	Gln	Ala	Leu	Lys	Ile	Phe
307															125	
309	ccg	tcg	tcg	gca	ttt	ggt	ccg	ggc	tac	atc	acg	gca	ctg	aaa	gca	gta
310	Pro	Ser	Ser	Ala	Phe	Gly	Pro	Gly	Tyr	Ile	Ser	Ala	Leu	Lys	Ala	Val
311															140	
313	ctt	ccg	ccg	gat	gtt	ccg	cta	ttt	gcc	gtc	ggc	ggc	gtg	acg	ccg	gaa
314	Leu	Pro	Pro	Asp	Val	Pro	Leu	Phe	Ala	Val	Gly	Gly	Val	Thr	Pro	Glu
315															160	
317	aac	cta	gca	caa	tgg	att	aaa	gca	ggc	tgt	gtg	ggc	gca	gga	ttg	ggt
318	Asn	Leu	Ala	Gln	Trp	Ile	Lys	Ala	Gly	Cys	Val	Gly	Ala	Gly	Leu	Gly
319															175	
321	agc	gat	ctc	tat	cg	ggc	caa	tcc	gtt	gaa	cg	acc	gca	cag	cag	
322	Ser	Asp	Leu	Tyr	Arg	Ala	Gly	Gln	Ser	Val	Glu	Arg	Thr	Ala	Gln	Gln
323															190	
325	gct	gca	ttt	gtt	aat	gca	tat	cga	gag	gca	gtg	aaa	tga			618
326	Ala	Ala	Ala	Phe	Val	Asn	Ala	Tyr	Arg	Glu	Ala	Val	Lys			
327															205	
330	<210>	SEQ	ID	NO:	6											
331	<211>	LENGTH:	205													
332	<212>	TYPE:	PRT													
333	<213>	ORGANISM:	Salmonella	typhimurium												
335	<400>	SEQUENCE:	6													
337	Met	Gln	Trp	Gln	Thr	Asn	Leu	Pro	Leu	Ile	Ala	Ile	Leu	Arg	Gly	Ile
338	1						5				10				15	
341	Thr	Pro	Asp	Asp	Ala	Leu	Ala	His	Val	Gly	Ala	Val	Val	Asp	Ala	Gly
342															30	
345	Phe	Asp	Ala	Ile	Glu	Ile	Pro	Leu	Asn	Ser	Pro	Gln	Trp	Glu	Lys	Ser
346															45	
349	Ile	Ser	Ser	Val	Val	Lys	Ala	Tyr	Gly	Gly	Arg	Ala	Leu	Ile	Gly	Ala
350															60	
353	Gly	Thr	Val	Leu	Lys	Pro	Glu	Gln	Val	Asp	Gln	Leu	Ala	Gly	Met	Gly
354	65											70			80	
357	Cys	Lys	Leu	Ile	Val	Thr	Pro	Asn	Ile	Gln	Pro	Glu	Val	Ile	Arg	Arg
358															95	
361	Ala	Val	Ser	Tyr	Gly	Met	Thr	Val	Cys	Pro	Gly	Cys	Ala	Thr	Ala	Thr
362															110	
365	Glu	Ala	Phe	Ser	Ala	Leu	Asp	Ala	Gly	Ala	Gln	Ala	Leu	Lys	Ile	Phe
366															125	
369	Pro	Ser	Ser	Ala	Phe	Gly	Pro	Gly	Tyr	Ile	Ser	Ala	Leu	Lys	Ala	Val
370															140	

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 04/10/2006
PATENT APPLICATION: US/10/572,976 TIME: 10:44:53

Input Set : F:\6550-086POA SeqListing.ST25.txt
Output Set: N:\CRF4\04102006\J572976.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/572,976

DATE: 04/10/2006

TIME: 10:44:53

Input Set : F:\6550-086POA SeqListing.ST25.txt

Output Set: N:\CRF4\04102006\J572976.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date